



The Liberation of Echo: A New Hearing for Film Sound

William Johnson

Film Quarterly, Vol. 38, No. 4. (Summer, 1985), pp. 2-12.

Stable URL:

<http://links.jstor.org/sici?sici=0015-1386%28198522%2938%3A4%3C2%3ATLOEAN%3E2.0.CO%3B2-S>

Film Quarterly is currently published by University of California Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/ucal.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact support@jstor.org.

The Libération of Echo: A New Hearing for Film Sound

Sound has been an integral part of the film for well over half a century, yet most critics and theorists still pay it little more than lip service. Even those who do write perceptively about sound treat it as a secondary attribute, like color, and continue to equate analysis of the image with analysis of the film as a whole.¹

In recent years attempts have been made to analyze film sound in greater depth, notably in special issues of *Yale French Studies*² and of *Screen*³ (which I refer to respectively as YFS and SST). Yet even here, contributors either retain the traditional view of sound as appendage to the image or swing to the opposite extreme, making sound the puppeteer of the image. One rare exception to the general trend is Elisabeth Weis's study of Hitchcock's use of sound:⁴ although concerned with specifics rather than theory, Weis consistently attributes equal importance to the two channels.

In this article I hope to show that sound in the film is a full and equal partner of the image. The first step is to examine some common misconceptions about the role of sound and indicate why they have remained widespread for so long.

The two main aspects of the traditional view of film sound are what Rick Altman, in the introduction to YFS, refers to as the *historical fallacy* and the *ontological fallacy*. The former consists of regarding the image as primary because it preceded the advent of the sound track. The latter consists of regarding the image as the essence of the film and sound as a pollutant.

As many writers have recognized, the historical fallacy ignores the prevalence of sound during the so-called silent era. In addition to the standard accompaniment of live music, there were experiments with adding sound effects and even live dialogue.⁵ Moreover, the

prehistory of the cinema had included attempts to marry the recording of images with the recording of sounds: as Balász points out,⁶ if those attempts had been successful they would hardly have been rejected in favor of the image alone. After all, the forms of entertainment that influenced the nascent cinema—*theater, vaudeville and music hall*—all relied heavily on speech, music and sound effects.

It is a widely held assumption that film-makers took a long time to realize the potential of the sound track. Thus Mitry⁷ states: "Except for rare masterpieces . . . , for several years [after the advent of recorded sound] there were only filmed plays and musicals." Yet the transition period was in fact remarkably brief. In 1929, the first year in which Hollywood was fully committed to the sound track, film-makers explored many ways of combining sound and image. They freely used speech and sound effects in outdoor scenes, even in such an unlikely context as Millard Webb's *Glorifying the American Girl* (whose lengthy company picnic sequence challenges just about every stereotype of the early sound film). Speech and sound effects were also freely used with frequent scene changes, camera movement or unusual camera setups (as in Capra's *Flight*, where a dialogue in a hangar is filmed in high-angle long shot). There were also many flirtations with experimental uses of sound. Examples include: synchronizing a source with a "wrong" sound, like the pizzicato playing of fingers on lips in Elsie Janis's *Paramount on Parade*; extradiegetic address of speech to the camera, as in Lubitsch's *The Love Parade*; expressionistic variation of dynamics, as in the well-known "knife" scene in Hitchcock's *Blackmail*; enjambement of sound belonging diegetically to one scene through part of the following scene, as in Sternberg's *Thunderbolt*. By 1931, virtually

all the uses of sound that can be found in narrative films today had already been pioneered. In short, most film-makers adapted rapidly to the sound track—in my view, because it was not an accretion but a completion.

The ontological fallacy is no longer proclaimed so passionately as in the late twenties and early thirties, but it survives in two widespread assumptions: that sound in general plays only a marginal role in film and that synchronized speech in particular is frequently redundant. Since speech comes under fire for other reasons, I deal with it later.

On the question of marginality, a typical assertion is that of Sparshott:⁸ “It is the requirements of the visual image that call for the elaboration of equipment and the circumstances of display that are fundamental to cinema. Film sound has no distinctive qualities in itself, and can be meaningfully discussed only as an adjunct to the visual.” Even critics who show a greater interest in sound allow it hardly any wider scope than does Sparshott. Thus Raymond Bellour, in his analysis of *Gigi*,⁹ automatically rejects any possibility that the structure of this musical might depend significantly on the sound track: “The image is sovereign on the level of syntagmatic demarcation. . . .” Yet in Demy’s *The Umbrellas of Cherbourg*, to take a clear-cut example, the duration of the whole film and of many individual scenes, the tempo of many actions and gestures, and the season and time of day were all determined to a large extent by the prerecorded sound track. In any instance where the sound may have a predetermined duration or structure the image may be viewed as the adjunct. It is the ontological fallacy that leads critics to analyze films as if the image were determined first, independently of the sound.

The same assumption also underlies—and undermines—several of the contributions to YFS. Claudia Gorbman refers¹⁰ persistently to “film/music relationships” as if the music were not part of the film. Daniel Percheron declares¹¹ that “the opposition, sound ‘on’/sound ‘off’ . . . depends on the image, and consequently testifies to the image’s primacy.”

Yet this primacy exists only in the nomenclature, not in the film: it is just as accurate to consider the image “on” or “off” the sound as the reverse. Metz, who recognizes this particular bias, unintentionally promotes another as he examines¹² what appears to be our Western cultural valuation of “visual objects” over aural objects: “From a logical point of view, ‘buzzing’ is an object, an acoustic object in the same way that a tulip is a visual object. . . . As soon as it becomes a question of naming the concept of aural object itself, it is necessary to add to the word ‘object’ the epithet ‘aural,’ . . . while no precision is required for that which should logically be called ‘visual object’: we consider it self-evident that a banner is an object (with no adjective needed) but we hesitate over a hoot; it’s an infra-object, an object that is only aural.” What Metz fails to realize is that “tulip” and “banner” refer not to *visual* objects but to objects that may be perceived by various senses. Thus the visual equivalents of “buzzing” and “hoot” would include not “tulip” or “banner” but “flash,” “blur,” and “red.”

Altman sets up another specious antithesis¹³ by stating that image-without-sound and sound-without-image are not complementary and symmetrical because “the former is a perfectly common situation in nature (a person standing quietly) while the latter is an impossibility (sounds are always produced by something imageable). Thus the completion of the former paradigm depends on the object within the image, while the completion of the latter depends on the auditor (who must look around and *find* the source of the sound). Images call for no action on the part of the auditor.” Yet many sounds are *not* imageable in any practical sense, beginning with the “room presence” that stands for silence and continuing through sounds with vast or nebulous sources, such as wind or thunder, or hidden sources, as in most powered vehicles, phones, radios, etc. Moreover, sound-without-(visible)-image has its counterpart in images where the source is visible but the sound is suppressed, as when characters speak at a distance from the camera or behind glass; and then the image does call for action on the part of the auditor, as he or she tries to supply the missing sound. By contrast, there can be scenes in which the auditor is under no compulsion to find the

As we go to press, Columbia University Press has just released an anthology of articles, *Film Sound: Theory and Practice*. William Johnson will review this volume for us in our following issue.

source of a sound "off." A classic example is the ending of Robson's *The Seventh Victim* when the sound of an overturned chair behind a closed door gives the auditor all the necessary information—that a character has hanged herself.

If sound does have equal ontological status with the image, it may be wondered why the assumption of image primacy is so widespread. Sparshott alludes to one reason in the statement cited earlier: "the circumstances of display that are fundamental to the cinema." In other words, the layout of the movie theater, especially the seating, is determined by the image. But this would imply primacy only if it conflicted with aural needs, which it does not, since the nature of sound propagation and perception imposes few constraints on the seating. In any case, the typical movie theater follows the same basic seating pattern as the legitimate theater, the music hall and the concert hall, in all of which sound is paramount.

A second reason for the assumption of image primacy is that full acceptance of sound would seem to validate the occasional attempts of other sensory channels (smell and touch) to enter the film process. Although such attempts have led nowhere, couldn't technological advances one day bring us smell-on-film and touch-on-film? Don't arguments against these other senses echo the pronouncements of many twenties critics that sound would be a passing fad?

There are nevertheless good reasons for believing that sound and image belong together in the film while the other senses will never play more than a novelty role. Taste and the various senses of touch are localized in or on the body, so that in any realistic film the "taster/feeler" would have to be identified with one character at a time (a problem that Aldous Huxley glosses over in describing the feelies of *Brave New World*). Smell is the only other sense that can consistently receive information from a distance, like vision and hearing, but its chemical means of transmission makes it slow to take effect and cumbersome to control. (Direct neuronal stimulation would bypass these difficulties only to raise greater ones.) Psychophysically, vision and hearing carry far more information than the other senses: the channel capacity is 40 bits

per second for the eyes and 30 bits per second for the ears; third place goes to the skin with only 5 bits per second, while taste and smell are estimated at 1 bit per second each.¹⁴ Image and sound can be made to carry information rapidly and precisely to viewers/auditors who, without special training, can instantly identify a particular face or voice—while most of us cannot distinguish a tulip from a rose with our nostrils alone or a dime from a penny with our fingers.

Valuation of the image over the sound track was reinforced by the circumstances of the latter's birth. Whereas the image had begun as a single strip of film, and could therefore be thought of as an entity even after it became a complex tissue of shots, the sound track started out as an obvious assemblage of diverse sources. The production of film music already existed as a separate enterprise; speech became a prime concern of actors and director; and sound effects fell by default to other specialists. So today, while analysts do not question the unitary status of the image (the *Grande Syntagmatique*, for example, was designed to cover the entire image strip, no matter what variations there might be in content, camera position, lighting or montage), they invariably attack the sound track by dividing it into three virtually autonomous parts—speech, music and sound effects. Thus Stephenson and Debrix¹⁵ can assert, "There are not different kinds of image as there are different kinds of sound—music, speech and noises," forgetting that the image strip may include stock shots, second unit footage, model shots and other special effects, and additional shooting after the first cut. The completed sound track is no more heterogeneous than the image strip.

Even though speech, music and sound effects may be produced separately, they clearly have no functional autonomy. Thus speech can function like sound effects, stressing its concrete origin—a distinctive timber of the voice, or a modifying channel such as a phone, a radio or (as in the Gregoretti episode of *RoGoPaG*) a buzzing throat mike. Speech can also function like music, stressing a rhythmic or intonational pattern, as in the chanting of a crowd or the newsreel narration in *Citizen Kane*, not to mention such experimental uses of rhythmic speech as in Paul

Sharits's *T,O,U,C,H,I,N,G*. Sound effects can function like music, as in the rhythmic hamburger-slapping of the group in Sayles's *Return of the Secaucus Seven*, or most obviously in the opening of Mamoulian's *Love Me Tonight* with its cumulative symphony of hammer blows, pickax crashes, and snores. Sounds can also convey information in place of language, as with coded knockings and horn blowings or with the cultural associations of sirens, whistling, finger tapping, and the like. Music can function like sound effects, either mimetically as in Grémillon's *Remorques*, where choral music is blended with storm noises, or as a diegetic background (the apparent product of a radio, stereo, street band, etc.); and it can be made to carry coded information, not only through such familiar "words" as the Wedding March and the Marseillaise but also by more individual means, such as the party music that continues faintly through a shot of lonely Sylvia Sidney in Sternberg's *An American Tragedy*.

In spite of this continual overlap of functions, many writers still condemn speech as inherently redundant to the image while accepting music and sound effects. Yet just how much is redundant when we see and hear a character speaking? Even if we can lipread and the character remains in full face, we cannot tell voice quality or intonation without hearing the spoken words. Besides, synchronized speech serves other purposes than the conveying of immediate information. In Hitchcock's *Vertigo*, for example, the deliberate dialogue scenes between James Stewart and Barbara Bel Geddes throw Stewart's growing obsession with Kim Novak into sharp relief.

In recent years, Marxist and Marxist-influenced analysts have reinforced traditional prejudices against film speech by focusing on what they consider its ideological role in capitalist societies, especially the US. In this view, redundancy of sound and image is essential to bourgeois films because it creates an illusion of unity that can assure the viewer/auditor of his/her own unity within the dominant ideology. Thus Stephen Heath asserts:¹⁶ "The stress is everywhere [in the classical cinema] on the unity of sound and image and the voice is the point of that unity: at once

subservient to the images and entirely dominant in the dramatic space it opens in them. . . ." Mary Ann Doane goes so far as to say¹⁷ that dialogue is privileged on the sound track in order to "preserve the status of speech as a property right. . . ."

However, the ideological case against synchronous sound rests on shaky evidence. Marxist critics invariably compare the use of sound in the commercial cinema with that in the films of Godard, Straub-Huillet, Oshima and others who work in critical opposition to, though still within, capitalist bourgeois societies. Yet the great majority of films made in other parts of the world display the same traits (including synchronized, privileged speech) that allegedly stigmatize the bourgeois cinema. This is true even of a sophisticated, consciously Marxist film such as Gutiérrez Alea's *Up to a Certain Point*.¹⁸

In any case, the prevalence of synchronized speech in the commercial film has been exaggerated. In most films, as speaking characters move, turn their heads, are cut "off," etc., synchronization can be identified only intermittently; when sync speech is sustained, it usually challenges the realistic conventions, as in two examples in which Spencer Tracy addresses the camera directly: Cukor's *Edward, My Son* and Minnelli's *Father of the Bride*. Not surprisingly, the most extensive use of sync speech is found in "art" films, such as Malle's *My Dinner with André* and Straub-Huillet's *Class Relations*, and above all in experimental films—Hodgdon's *The Whole Film, A Filmic Relationship* and *A Prepared Text*, Greenaway's *Act of God*, Landow's *Wide Angle Saxon*, and much of Snow's rambling exploration of image-sound relations, *Rameau's Nephew*.

Tom Levin is quite correct when he refers¹⁹ to the "specificity of the acoustic and the visual domains" and states that "the closer word and image are coupled the greater the contrast between them becomes manifest." But his conclusion—that this contrast poses a threat to the realistic narrative cinema because it is "an accurate symptom of a culture whose members are alienated from themselves"—would, if true, apply to every film-producing society in the world. In fact, the contrast between sound and image has been

a continuing source of strength in all kinds of films. Sound and image have remained independent not only covertly, by virtue of their separate recording processes, but also to a large extent overtly, in their relations within the finished film. The rest of this article examines the scope of those relations.

Sound and Image Compared

The crucial condition of sound and image in film is that they are both *distinct* and *comparable*.

Points of distinction. Because the image is usually taken as the norm from which sound deviates, differences between the two channels tend to be exaggerated. To deal briefly with some common assumptions: (1) *The image is sharply bounded while sound is not.* But the microphone records a limited field of sound much as the lens records a visual field, even though the two rarely coincide. In any case, reflections, shadows and halation can challenge the sharp boundary of the screen image. (2) *The image is spatial while sound is temporal.* But the image can assume a temporal quality when it is systematically displaced, either by rapid cutting or by camera/lens movement; and sound can indicate spatial shifts when it changes with or within scenes. (3) *The image is susceptible to frequent change by cutting while sound is not* (in a realistic context, at least). But this assumes that cutting is the most important means of film segmentation, and even for the image it is hardly self-evident that a major change during a continuous scene—an explosion of movement, the sudden entry of a character, the switching on or off of lights—marks less of a break than a cut. Fred Mogubgub's *Enter Hamlet* (in which a different picture accompanies each word of the "To be or not to be" soliloquy) serves as a reminder that a continuous but articulated sound may convey the same effect of segmentation as a strip of cut images.

The one important distinction between sound and image stems from their physical origin. The majority of objects that are represented in the photographic image do not generate light but reflect it (and on the screen, the entire image consists of reflected light). By contrast, virtually all objects that emit sounds do so actively, through some kind of

movement, and this remains true of the theater loudspeakers that reproduce those sounds. Thus sound tends to be a series of active events that is *experienced* while the image tends to be a static display that is *read*. This distinction enables the two channels to complement each other, carrying large amounts of information without mutual interference.²⁰

Points of comparison. These embrace both primary codes—formal or cinematic—and secondary codes—cultural or extracinematic.

At the formal end of the range, sound and image may share the extreme similarity of being either present or absent. The absence is rarely absolute, however: for the image, it consists of the perceptible grayness of the "black" or imageless screen, and for the sound, of a faint background hum. Thus in each channel the *presence* of absence constitutes a ground of comparison with the other.

Sound and image can also be compared in terms of such basic formal parameters as duration, intensity, multiplicity and rate of change. Some of these are directly comparable, such as the duration of an image (event) and of a sound (event), or the multiplicity of elements in each. Others involve synaesthetic equivalences, as between the intensity of the image (brightness) and of the sound (a combination of pitch and loudness). Even in these latter cases, however, the relation is not merely subjective but derives from physically measurable conditions.

As Alan Williams notes briefly in YFS,²¹ both sound and image are subject to virtually the same set of manipulations. Just as the camera can be placed at varying distances from the subject, or moved bodily in the process of recording the image, so can the microphone; and the use of unidirectional and omnidirectional mikes corresponds to the use of long- and short-focus lenses. Punctuation such as the cut, fade and dissolve is as familiar on the sound track as in the image strip. All of the most widely used manipulations—both modal and segmental—are directly comparable between the two channels.²²

At the cultural end of the range of similarities, both sound and image may be derived from profilmic or synthetic material. In the former case, the material used in either channel may be staged or unstaged, diegetic or

extradiegetic, subjective or objective in point of view/hearing.

With all of the foregoing parameters sound and image may vary independently, even in realistic narrative films: thus at any given moment the two channels will be perceptibly similar or dissimilar in at least one of many ways. For example, sound may be present while the image is absent (the screen is black) as at the end of LeRoy's *I Am a Fugitive from a Chain Gang*, when Paul Muni fades into the darkness with the words "I steal"; the image may be present while sound is absent (the sound track is silent), as in Joseph Lewis's *The Big Combo* when Brian Donlevy's hearing aid is pulled out (a scene that also illustrates the combination of objective image and subjective sound); and of course, as more usually occurs, sound and image may be co-present.

Special relationships. Discussion of sound in films often concentrates on two special relationships between sound and image: the temporal link of *synchronization* and the spatial link of so-called "sound on" and "sound off," for which I prefer the unbiased terms *conjunction* and *disjunction*. However, these two relationships are ancillary to the points of comparison outlined above.

Although synchronized speech is usually treated as a separate category, the two special relationships obtain with all kinds of sound. Their main effect is to qualify the basic similarities and dissimilarities between sound and image. The weakest case is conjunction, which may be little more than the neutral bracketing of the two channels. The other three cases—disjunction, synchronization, and what may be called anti-synchronicity (the deliberate avoidance or disruption of sync)—can all add emphasis to either a similarity or a dissimilarity. The strength of the emphasis will depend on the context, varying inversely with the frequency or expectedness of the case. In *The Love Parade*, for example, the cutting of scenes in time with rhythmic noises is an infrequent use of sync that strongly emphasizes the temporal similarity of sound and image.

Sound and Image Together

Two important questions remain: what do the sound-image relations signify, and how

does meaning emerge from their abundance?

The answers that follow are tentative, being based on a rudimentary structural system of interactions between the two channels. As simple as this system is, however, it does help to resolve many problems arising from the assumption of one dominant channel in the film.

Two caveats. First, the system is not meant to *replace* the perception of meaning in the individual channels: instead, it qualifies and supplements that meaning. Second, in order to test the general validity of the system, I cite examples ranging from commercial entertainments to serious political and experimental films. The fact that my focus is what these films have in common is not meant to deny their differences.

In each of the main group of relations (points of comparison) there are at any given moment two basic possibilities: sound and image may coincide or differ. The former case may be termed *confirmation*; the latter, *opposition*. Here are some examples of the basic paradigms:

Confirmation. In Bergman's *The Seventh Seal*, Max Von Sydow as the knight relaxes with Nils Poppe and Bibi Andersson in an impromptu picnic. Von Sydow refers to the mood of unexpected serenity by saying quietly, "I shall remember this moment. . . ." As he raises a bowl of milk to drink, reflections from its surface cause his face to literally light up. Both the sense and acoustic quality of the speech combine with the double reflection of the image in a vivid demonstration of the mood.

In Lester's *A Hard Day's Night*, the Beatles run down a fire escape to the accompaniment of "Can't Buy Me Love." The camera, placed below, rotates to follow the singers as they zigzag from flight to flight; at the same time, sunlight burns radiant diamond-shaped patterns in the openwork of the steps. As the song reaches the line "I'll buy you a diamond ring, my friend," the conjunction of the words "diamond" and "ring" with the diamond-shaped patterns and the circular movement of the image creates a sense of euphoric fitness.

At the end of Meszaros's *Diary for My Children*, set during the post-World War II

Stalinist era in Hungary, the young protagonist goes to visit her jailed mentor. They have to speak through two separate chain-link fences in an enclosed yard. Her last words are "You've turned gray"; after that the image freezes, enabling the viewer to take stock of the grayness of the scene, and then fades to a uniform gray. Aided by the additional segmentation of the image, the remark expands from a physical description of a man's hair to a metaphor for a whole society.²³

In Ashby's *Being There*, Peter Sellers plays a gardener who has never known any life outside his wealthy employer's town house. Then the employer dies. As Sellers sits on the bed gazing at the corpse, there is a repeated disjunctive noise—the whining attempt to start an automobile engine. Although the sound serves a naturalistic purpose, its disjunction also imposes a relation with the silent focus of the image, Sellers confronting the corpse: subjectively, it suggests the wish that he could restart the vital functions of his employer; objectively, it foreshadows the difficulty he faces in making a fresh start in life.

Opposition. In *Citizen Kane*, there is a strong contrast of intensities when Welles as Kane dies: a big close-up of his mouth accompanies his almost whispered "Rosebud"—an opposition intensified by the obvious synchronization. The ending of Szabo's *Mephisto* reverses the contrast: Klaus Maria Brandauer, as the non-Nazi actor who has gone along with the Nazis in post-1933 Germany, is seen in long shot as a small spotlighted figure in a vast arena, while the deep voice of Rolf Hoppe (as a Nazi general whom the actor thought he could influence) booms at him over loudspeakers. Here, the opposition is intensified by its conformity to the physically active nature of sound and passive nature of the image. In each of these two examples, the unusual combination of sound and image draws attention to a pivotal development—a clue to Kane's character, the exposure of the actor's delusion.

A different pair of contrasts is found in Lang's *Fury*. At the trial of the lynch mob, silent newsreel footage is projected as evidence for the prosecution. The silence heightens the abnormality of the defendants' ex-

pressions of hate and glee as they set fire to the jail and sabotage attempts to save it. The presumed victim, Spencer Tracy, is alive but in hiding, bent on revenge. When his brothers and fiancée urge him to reveal the truth he yells, "I don't need other people!" and stalks out into the night. Then, feeling alone in the deserted streets, he enters a bar from which music is blaring—only to find it empty except for a bartender, who stands for a moment like a statue. Here, the combination of loud sound and virtually immobile image heightens the sense of aloneness.

Three points emerge from these examples. The first point concerns the role played by spoken (or sung) language in sound-image relations. In the examples from *The Seventh Seal*, *Citizen Kane* and *Mephisto*, the relation depends partly on the acoustic properties of speech: while the sense of what Von Sydow and Hoppe say contributes to the relation, the words they use could be paraphrased without causing any significant change. With *A Hard Day's Night* and *Diary for My Children*, however, the relation hinges on a precise choice of words. As far as sound-image relations are concerned, there is no essential difference between meaningful speech and any other kind of sound.

Second, it is clear that image-based analysis would by no means class all of the relations as paradigmatic, since they range in duration from a brief part of a scene (*Kane*) to two or more fairly lengthy scenes (both *Fury* examples, *Mephisto*). On the basis of sound-image relations, however, even the lengthiest examples remain unitary, for one of two different reasons. The dominant relation may simply be repeated, as in the *Fury* trial, *Mephisto* and also the single but lengthy scene in *Being There*. Alternatively, the relation may be founded on—and measured by—a continuous sound passage rather than an image scene, as in the *Fury* bar episode with its continuous blare of music. It is important to analyze a film on the basis of sound-image relations rather than by image segments alone.

The third point is that some of the foregoing examples involve additional sound-image relations. Thus *A Hard Day's Night* contains the more general confirmation of vigorous image movement and vigorous music;

the quiet background in *The Seventh Seal* matches the near-stillness of the image; and in *Diary for My Children*, the metronomic clack of a pacing guard's boots echoes the geometric division of the image space by chain-link fences. While each of these supplementary relations happens to support the dominant relation, the reverse is equally possible: thus in *Mephisto* there is an element of confirmation between Hoppe's loud voice and the size of his body in the foreground of the image. In view of the multiplicity of sound/image parameters, few synchronic relations are likely to be simple. How can such complex relations be assessed?

Provisionally, it seems that there is an additive effect (with counter-examples assigned a negative value). In other words, supplementary relations of the same sign increase the strength of the confirmation or opposition, while those of opposite sign weaken it. The total effect varies with the strength of the individual relations, especially the dominant one. In what follows, however, I distinguish only two levels of intensity for each basic relation—strong and weak. Strong implies that the dominant relation is intense and/or is supported by supplementary relations of the same sign, with few or no counter-examples. Weak implies that the dominant relation is nonintense and/or is supported by few or no relations of the same sign and/or is accompanied by one or more significant counter-examples.

There are also, of course, sound-image units in which the sum is close to zero, either because there is no marked confirmation/opposition or else because relations of both types are numerous. (An example of the former might be a long shot of an empty landscape accompanied by fairly bland music; of the latter, the segment near the opening of *Rear Window* in which the voice of James Stewart's editor is heard on the phone, together with diegetic sounds, while the camera roams around the room and the courtyard.) The meaning of such units depends largely on the context. In fact, even the strong units cannot be fully appraised in isolation but must be related to their syntagmatic organization.

Sound-image syntagmas. Since sound-image units can vary greatly in length, and since the segmentation of sound often does not coincide

with that of the image, it may seem difficult to determine where one unit ends and another begins. In the *Seventh Seal* example, since the milk-drinking obviously does not coincide exactly with Von Sydow's words, and since he continues speaking for some time, where does the cited unit of strong confirmation yield to another, weakly confirmative unit?

As with the summing of individual units, precision seems unnecessary. In the foregoing example there is a continuing balance of confirmation with no significant change in sound or image. The relation established by one unit remains in force until the advent either of a different relation or of the same relation expressed with a marked difference. Determining just how many basic units there are in this higher-level segment is unimportant—at least at this stage.

Although the beginning or end of a traditional sequence may coincide with a significant change in sound-image relations, sequences based on the image alone can diverge so widely from those based on both channels together that I will refer to the latter as *passages*. A good example of the divergence between sequences and passages is provided by Conway's *Red Headed Woman*, a crypto-feminist film whose effectiveness is hard to account for in image-based terms. The film is a star vehicle for Jean Harlow, and since the outstanding element in her performance is her voice, it is easy to dismiss the direction of the image as tributary and inferior. Yet the interaction between sound and image organizes the apparently rambling plot into a strong chain of passages. At the beginning of the film, for example, Harlow attempts to seduce her boss Chester Morris and is frustrated by the return of his wife Leila Hyams. Harlow's uncerecermonious departure from Morris's house is followed by a close-up of her laughing. The combination of large image and loud laughter, the strongest confirmation in the film so far, proclaims the end of one passage and prepares the viewer/auditor for another. Later, after Morris and Hyams have divorced, Hyams visits him in hopes of a reconciliation only to find that he has just married Harlow: in a tony voice, Hyams delivers an intense wronged-woman speech and makes a proud exit, at which point Harlow regains control

by turning up the volume on her perky whine and uttering just three unexpected words: "That cheap thing!" Here it is a strong opposition that marks the end of the passage: the contrast between Harlow's words and tone of voice is reflected and reinforced by her petulant expression and arms-akimbo pose.

There can be interactions on a larger scale between different sound-image passages. The mood of serenity in the *Seventh Seal* interlude derives partly from its contrast with two preceding passages—the procession of flagellants and the attack on Nils Poppe at the inn. Both represent strong confirmation of high intensity: loud sound and continual image movement (plus, in the procession, a multiplicity of image elements). The euphoric scene in *A Hard Day's Night* stands out by contrast with a more static preceding passage, in the same way as traditional musicals intensify the strong confirmation of song plus dance by means of opposition or weak confirmation in the surrounding nonmusical passages.

Above the sound-image passage there does not usually seem to be any significant level of organization other than the *whole film*. Locating the boundaries of this structural unit causes few problems, of course. But the relation between the passage and the whole can be more elusive. It can be analyzed most easily in experimental films that have little or no narrative detail. Dwoskin's *Jesus' Blood Never Failed Me Yet* consists of many repetitions of a single image scene (a man walking toward the camera) accompanied by a male voice singing the evangelical hymn of the title. After a while soft chords from a chamber orchestra enter behind the voice and become progressively fuller in sound as they continue from repetition to repetition. The passage shift from weak confirmation (fairly close image of man, fairly loud voice) to weak opposition (distant image, fairly loud voice) at the end of each repetition fades in importance beside the gradual shift in sound-image relations that extends through the whole film. Here the latter shift is easy to identify because the image is repeated. Otherwise, an analysis based primarily on the image might not go beyond the former.

Because sound and image share nearly all of the same parameters and manipulations

they can take either side of any particular relation and, indeed, frequently exchange sides in a continuing or repeated relation. A case in point is the use of the voice-over narrator. Brian Henderson states²⁴ that in classical films such narrators are "jerked on and off stage in a manner that is quite undignified" and are "ludicrous stand-ins for the novelistic 'I,'" yet adds: "This is not to disparage the convention of the voice-over in cinema, which has figured in so many excellent films." Earlier Henderson hints at a resolution of this apparent contradiction: "Of course, voice-over narration in cinema does not comprise the whole text [as character narration can in literature]. . . . It is just one element among many elements, to be juggled along with them, often in shifting combinations." It is precisely as one sound element among the shifting combinations of sound and image that voice-over narration may be "jerked on and off stage" and yet form an integral part of the film.

Exchanges between the two channels affect not just voice-over narrations but all kinds of sounds (and images) as well. A sound-image relation may persist even though the individual parameters of the sound and the image change. For example, if lengthy images accompanied by brief sounds give way to brief images accompanied by lengthy sounds, there need be only a superficial change: the basic long/short opposition continues. This frequently occurs in Ozu's later films, where sequences involving characters consist of lengthy image scenes accompanied by rapid and/or brief passages of dialogue, while transition sequences consist of brief image scenes (streets, alleys, buildings, corridors, etc.) accompanied by continuous music. In this way Ozu achieves variety without disrupting the basic rhythm of his films.

The fact that sound-image syntagmas may have meaning over and above that of the individual relations is demonstrated by Naruse's *Meshi*. The film revolves around Setsuko Hara, a housewife who becomes aware of boredom after she and her husband, Ken Uehara, move from Tokyo to Osaka; she returns to Tokyo on her own, intending to take a job, but finally decides to stay with her husband. The first half of the film is accom-

panied by one of Fumio Hayasaka's typical westernized background scores, a vague, sentimental wash of music. Thus there is a continual weak opposition that happens to fit Hara's passive discontent. In the second half of the film, after Hara decides to leave for Tokyo, the music is often absent, and when it does occur it is more closely matched to the mood—a shift to a continual weak confirmation that gently underscores Hara's shift to action. *Meshi* demonstrates that even an insipid background score cannot be automatically condemned, as it usually is, for aesthetic or ideological shortcomings.

One obvious challenge to my proposed system is raised by films whose sound-image units are in such continual flux that no sustained or dominant relation emerges at higher levels. Yet this flux itself may constitute a dominant pattern—a series of contrasts that can be used to shock the filmgoer to either comic or serious effect. The Marx Brothers films offer examples of the former, with the brothers themselves displaying a wide range of sound-image relations: Groucho uttering wisecracks as he moves in his crouching lope, or rising to full height to launch a denunciation or a song; Chico's air of innocent reasonableness as he proposes ridiculous schemes in mangled English, or his cheerful piano playing with index finger pointed like a gun; Harpo's muteness accompanied with maniacal enthusiasm, lechery, misery or rage, at peace only when he plays the harp.

Both humor and seriousness emerge from the flux of such experimental films as Berliner's *Myth in the Electric Age* and *Natural History*, in which a loosely associated series of images is paired with independently chosen sounds. Some of the pairings are incongruous or ironic, such as a sawing noise with the image of a woodpecker or piano music synchronized with a man hammering, while others produce a strong confirmation, such as the dramatic jet roar that accompanies a time-lapse flow of mist down a mountain valley.

A serious application of sound-image flux is found in the early sound films of Vertov, *Enthusiasm* and *Three Songs About Lenin*. (*Enthusiasm* has the distinction of being the first film in which the sound-recordist is actu-

ally visible, just like Vertov's *Man with the Movie Camera*.) Both image and sound tend to be cut rapidly, though sometimes sounds continue through several image scenes and, less often, image scenes continue through several sound phrases. Frequent superimposed titles add complexity to the image, often changing or repeating in a separate rhythm of their own. A single held musical tone may accompany a crowd of workers (*Enthusiasm*) or the roar of an urban crowd may accompany a rural scene (*Lenin*). Images range from big close-ups to extreme long shots, and undergo such evident manipulations as freezing, time lapse and prismatic replication. The sound track includes synchronized speech, voice-over, disjunctive noises and various styles of music. Except for some anti-religious scenes near the beginning of *Enthusiasm*, the two films attempt no polemic—in fact, their mercurial sound-image relations could hardly sustain one. What they do offer is an impression of vigor and diversity that is meant to characterize the Soviet state.

The conspicuous role that sound plays in Vertov—as in such films as the Tavianis' *Padre, Padrone* and Godard's *First Name: Carmen*—is only one aspect of liberated Echo. Since my thesis is that sound deserves equal billing with the image even when its role is inconspicuous, I choose a more naturalistic film to make a final point.

Critical response to Wenders's *Paris, Texas* has differed most sharply over the peepshow palace episode, which is seen as either a high point or a letdown. Those in favor of the episode could invoke sound-image relations to support their view: an exchange of roles between the two channels enables the film to end both differently from and in balance with its beginning, since Harry Dean Stanton evolves from one sound-image opposition—mobile and mute—to its inverse—static and voluble. However, dissenting critics also could invoke sound-image relations: the peepshow episode does not actually end the film, and its dialogue, full of tardy exposition, lacks the driving intensity of Stanton's walking.

In short, the sound-image system I have outlined is not meant to provide easy answers. In more positive terms, this theoretical structure should not constrain discussion about

films or foreclose on critical options but, rather, open up new perspectives on the film.

NOTES

1. As one of many examples, Jurij Lotman, in *Semiotics of Cinema* (tr. Mark E. Suino, Michigan Slavic Contributions, Ann Arbor, 1976) begins by acknowledging the long history of film sound, but later asserts: "The pictorial language . . . is dominant. If we peel off the photographic level from a . . . complete film, we are doing much the same thing as a linguist who studies . . . speech activity in general. On a certain level this approach is not only possible, but even necessary."
2. "Cinema/Sound," Yale French Studies No. 60, 1980 (YFS).
3. "On the Soundtrack," *Screen XXV* 3, May-June 1984 (SST).
4. Elisabeth Weis, *The Silent Scream: Alfred Hitchcock's Sound Track*. London & Toronto: Associated University Presses, 1982.
5. See for example Harry M. Geduld, *The Birth of the Talkies*. Bloomington and London: Indiana University Press, 1975.
6. Béla Belázy, *Theory of the Film*. New York: Dover, 1970.
7. Jean Mitry, *Esthétique et Psychologie du Cinéma*, Vol. 2. Paris: Editions Universitaires, 1965.
8. F. E. Sparshott, "Basic Film Aesthetics," *Journal of Aesthetic Education*, April 1971: cited in *Film Theory and Criticism: Introductory Readings*, ed. Gerald Mast and Marshall Cohen. New York: Oxford, 1974.
9. Raymond Bellour, "Segmenting/Analysing," *Quarterly Review of Film Studies*, 1, 3, August 1976; published in a new translation by Diana Matias in *Genre: The Musical*, ed. Rick Altman. London/Boston/Henley: Routledge & Kegan Paul, 1981.
10. Claudia Gorbman, "Narrative Film Music," YFS.
11. Daniel Percheron, "Sound in Cinema and Its Relation to Image and Diegesis," *Ça/Cinéma* No. 2, October 1973, reprinted in YFS.
12. Christian Metz, "Aural Objects," from "Le Perçu et le Nommé," in *Pour une esthétique sans entrave—Mélanges Mikel Dufrenne*. Paris: Editions 10/18, 1975, reprinted in YFS.

13. Rick Altman, "Moving Lips: Cinema As Ventriloquism," YFS.
14. *Fundamentals of Sensory Physiology*, ed. Robert F. Schmidt. New York/Heidelberg/Berlin: Springer-Verlag, 1978.
15. Ralph Stephenson and Jean R. Debrix, *The Cinema As Art*. Harmondsworth: Penguin, 1965.
16. Stephen Heath, "Narrative Space," in *Questions of Cinema*, Bloomington: Indiana University, 1981.
17. Mary Ann Doane, "Ideology and the Practice of Sound Editing and Mixing," in *The Cinematic Apparatus*, ed. Teresa de Lauretis and Stephen Heath. New York: St. Martin's, 1980.
18. The universality of "bourgeois" practices does not of course prove that they are innocuous, but it does throw doubt on analyses that take their ideological thrust for granted.
19. Tom Levin, "The Acoustic Dimension: Notes on Cinema Sound," SST.
20. Sound and image can of course be made to interfere with each other. In Fassbinder's *In a Year of Thirteen Moons* many auditors are probably distracted from Volker Spengler's reminiscences in a slaughterhouse by the sight of mechanized steer-killing in the background. More often the two channels interact synergistically. In Larry Gottheim's *Mouches Volantes*, which permutes a set of visual sequences alternately with silence and with an unrelated conversation, each sequence with sound acquires a greater apparent spatial depth than its silent counterpart.
21. Alan Williams, "Is Sound Recording Like a Language?," YFS.
22. Tom Levin, op. cit., is in error when he states that "there is no acoustic equivalent of the freeze frame." The frame is not literally frozen but repeated, a technique that can easily be applied to sound. An example occurs at the end of Bellochio's *Fists in the Pocket* when Lou Castel falls into a fit and dies: as the image freezes, the high note of an operatic aria that is playing on a phonograph is also "frozen"—that is, prolonged by looping.
23. I am indebted to Yvette Biro for confirming that the original Hungarian phrase for "to turn gray" has similar connotations to its English translation.
24. Brian Henderson, "Tense, Mood, and Voice in the Film," *FQ XXXVI*, 4, Summer 1983.

Reviews

STOP MAKING SENSE

Featuring the Talking Heads. Director: Jonathan Demme. Producer: Gary Goetzman. Director of Photography: Jordan Cronenweth. A Cinecom International Films/Island Alive Release.

With astonishing swiftness, the rock 'n' roll documentary—the "rockumentary," if you will—developed a routine as predictable as the B western. D. A. Pennebaker's epochal *Don't Look Back* (1967) and *Monterey Pop* (1968) are the obvious prototypes, their makeshift mix of concert sequences, *cinéma vérité* backstage passages, and color commentary by

semi-articulate scene-makers inspiring virtual xeroxes from enterprising film-maker/groupies desperate for screen subject matter. Early on, the form attained a creative summit with the yin/yang dialectics of *Woodstock* (1970) and *Gimme Shelter* (1971), but from there it was mostly downhill. Throughout the next decade, bands from Abba to Zappa put their acts on celluloid in what had become a cinematic analog to vanity publishing: Alice Cooper's *Welcome to My Nightmare*, AC/DC's *Let There Be Rock*, Blue Oyster Cult and Black Sabbath's *Black and Blue*, or Paul McCartney's *Rockshow*. Well before the Roll-

LINKED CITATIONS

- Page 1 of 1 -



You have printed the following article:

The Liberation of Echo: A New Hearing for Film Sound

William Johnson

Film Quarterly, Vol. 38, No. 4. (Summer, 1985), pp. 2-12.

Stable URL:

<http://links.jstor.org/sici?sici=0015-1386%28198522%2938%3A4%3C2%3ATLOEAN%3E2.0.CO%3B2-S>

This article references the following linked citations. If you are trying to access articles from an off-campus location, you may be required to first logon via your library web site to access JSTOR. Please visit your library's website or contact a librarian to learn about options for remote access to JSTOR.

Notes

⁸ **Basic Film Aesthetics**

F. E. Sparshott

Journal of Aesthetic Education, Vol. 5, No. 2, Special Issue: Film II, the Teaching of Film. (Apr., 1971), pp. 11-34.

Stable URL:

<http://links.jstor.org/sici?sici=0021-8510%28197104%295%3A2%3C11%3ABFA%3E2.0.CO%3B2-J>

²⁴ **Tense, Mood, and Voice in Film (Notes after Genette)**

Brian Henderson

Film Quarterly, Vol. 36, No. 4. (Summer, 1983), pp. 4-17.

Stable URL:

<http://links.jstor.org/sici?sici=0015-1386%28198322%2936%3A4%3C4%3ATMAVIF%3E2.0.CO%3B2-A>

NOTE: *The reference numbering from the original has been maintained in this citation list.*